Message

From: McGrath, Jesse [/O=EXCHANGELABS/OU=EXCHANGE ADMINISTRATIVE GROUP (FYDIBOHF23SPDLT)/CN=RECIPIENTS/CN=C34718C70BB64FE58E72FA1AB387E90E-JMCGRA02]
Sent: 3/9/2017 8:24:10 PM
To: Joseph Warfel [joseph.warfel@gmail.com]
Subject: RE: Do you have any good ways to explain independence of errors when taking measurements?

Sorry I didn't see this before the meeting. They might be biased, but most instruments have very little bias – the goal of the QC checks is to measure that bias, and respond to changes to keep it near 0%. In the chart below the instrument seems unbiased, but is very imprecise. Meaning they're probably careful about calibrating it, but not keeping up on other maintenance.

From our conversation today I view it this way in case you need to feed it on.

- The result of your measurements is the convolution of the distribution you want to measure, and the distribution of measurement errors from the instrument. Measurement errors being different from the error which is the difference between the sample mean and the population mean. In this case it's the difference between the "true" value of an observation and the value the instrument reports. These are usually independent over the time we sample them. I don't think that matters for what mistake people make, though.
- With a variation of the test of power you can find your required sample size, and how "bad" the measurement errors can be to control your false positive and false negative rates. Our goals in this case are a mean and standard deviation of 7% or less.
- Over time we measure the error distribution by measuring a "known" value and determining the difference from the known. The mean of that is the bias of the instrument, and standard deviation is the precision. We call the observations "QC Checks" here.
- Any observation that comes from the error distribution must be included, even if it's "bad", or you'll say the quality of the instrument is better than it really is. Any observation that does not come from that distribution must be removed. "Not coming from that distribution" means an instrument failure or other event that means you've lost control of the instrument. Since that flawed distribution is convolved with the population distribution that you measured at the time, those population values must also be removed.

What is being pushed recently is to remove every QC check outside of 7%, even if it seems to come from the error distribution and is not an outlier, and to remove the population data back to the last "good" QC check inside of 7%. Even if the check outside of 7% otherwise appears to come from the distribution. Some people here genuinely believe this improves the quality of their data, though it's clearly a case of mistaking the map for the territory, and is egregious misreporting of data. They do that because they misunderstand concepts in statistical quality control that Shewhart and Deming developed.

I'm looking for documentation, texts, cases studies, or contacts that can get those, to explain how bad that practice is. The audience effectively does not understand or misunderstands statistics, so statements like "you shall not do this" or "this person was punished/fired/fined/imprisoned for this" will be more effective than an explanation of the statistical basis of the mistake.

Thank you,
Jesse

From: Joseph Warfel [mailto:joseph.warfel@gmail.com]
Sent: Thursday, March 09, 2017 11:38 AM
To: McGrath, Jesse <mcgrath.jesse@epa.gov>
Subject: Re: Do you have any good ways to explain independence of errors when taking measurements?

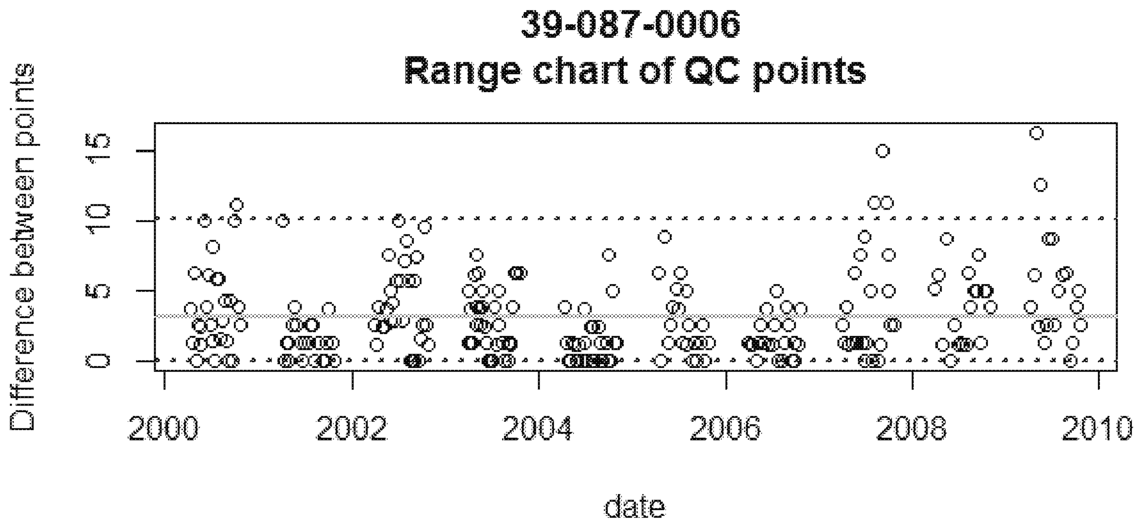
Hello Jesse,

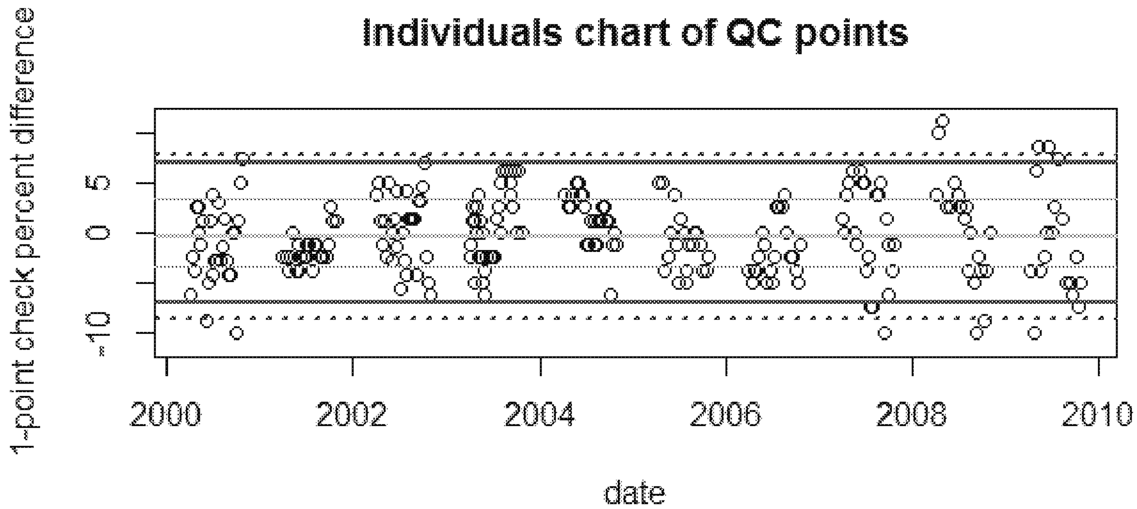
Are the measurement errors unbiased?

-Joseph

On Fri, Mar 3, 2017 at 5:02 PM, McGrath, Jesse <mcgrath.jesse@epa.gov> wrote:

> I'm trying to explain to people that should know better that you can't just invalidate data when a QC check exceeds an absolute limit because it's likely that you're just clipping the ends off your distribution and misrepresenting it, rather than actually invalidating "bad" data.
>
> This takes a bit to explain if you're more used to dealing with the distribution of the attribute of interest, rather than the distribution of the instrument's errors that's used to measure the attribute.
>
> There's a pic below of what I mean, you can ignore the top chart for now. The bottom chart is percent differences from a constant value over time.
>
> The instrument isn't that good and the standard deviation is pretty large. The standard deviation should be lower than 7%, and it's getting close to that limit. People keep trying to just "invalidate" every individual point outside of 7% (the purple lines) and say problem solved! But that doesn't actually fix anything.
>
> It's not particularly easy to explain, even to people who know what they're talking about and I'm wondering if you've got anything ways to present it.





Jesse McGrath

US EPA Region 5

Desk Phone: 312 886-1532

Email: mcgrath.jesse@epa.gov